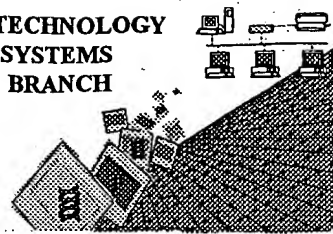


## **RAW SEQUENCE LISTING** **ERROR REPORT**

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/786,880B  
Source: Pg/09  
Date Processed by STIC: 4/12/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER**  
**VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND  
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efb/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002



PCT09

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/786,880B

DATE: 04/12/2002

TIME: 14:58:45

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

Does Not Comply  
Corrected Diskette Needed

3 <110> APPLICANT: AVENTIS PHARMA S.A.  
5 <120> TITLE OF INVENTION: ESSENTIAL GENES FROM C. ALBICANS AND A METHOD FOR  
6 SCREENING ANTIMYCOTIC SUBSTANCES USING SAID GENES  
8 <130> FILE REFERENCE: 16655  
10 <140> CURRENT APPLICATION NUMBER: US/09/786,880B  
11 <141> CURRENT FILING DATE: 2001-03-08  
13 <150> PRIOR APPLICATION NUMBER: EP98402255.8  
14 <151> PRIOR FILING DATE: 1998-09-11  
16 <160> NUMBER OF SEQ ID NOS: 25  
18 <170> SOFTWARE: PatentIn Ver. 2.1

see pp 4-6, 8-11

## ERRORED SEQUENCES

1529 <210> SEQ ID NO: 11  
1530 <211> LENGTH: 1847  
1531 <212> TYPE: PRT  
1532 <213> ORGANISM: Candida albicans  
1534 <400> SEQUENCE: 11  
1535 Met Ser Gly Ile Phe Asn Trp Ser Leu Asp Val Phe Ala Asp Ile Tyr  
1536 1 5 10 15  
1538 Asn Thr Leu Lys Phe Glu Ser Asn Ile Asp Leu Asp Thr Ile Asp Phe  
1539 20 25 30  
1541 Thr Ser Ile Lys Asn Asp Leu Ala Asn Val Leu Ile Thr Pro Val Pro  
1542 35 40 45  
1544 Leu Asp Gln Ser Arg Ser Lys Leu Gly Asp Ala Ser Lys Pro Val Ala  
1545 50 55 60  
1547 Leu Pro Ser Gly Asp Glu Val Lys Leu Asn Gln Ala Ser Ile Glu Ile  
1548 65 70 75 80  
1550 Thr Gly Val Leu Ser Asn Glu Leu Asp Leu Asp Glu Leu Asn Thr Ala  
1551 85 90 95  
1553 Glu Leu Leu Tyr Asn Ala Ser Asp Leu Ser Tyr Lys Lys Gly Thr Ser  
1554 100 105 110  
1556 Ile Gly Asp Ser Ala Arg Leu Ala Tyr Tyr Leu Arg Ala His Tyr Ile  
1557 115 120 125  
1559 Leu Asn Ile Val Gly Tyr Leu Val Ser His Lys Arg Leu Asp Ile Ile  
1560 130 135 140  
1562 Thr Asn Asn Asn Gln Val Leu Phe Asp Asn Ile Leu Lys Ser Phe Ser  
1563 145 150 155 160  
1565 Lys Ile Tyr Thr Leu Ser Gly Lys Leu Asn Asp Met Ile Asp Lys Gln  
1566 165 170 175  
1568 Lys Val Thr Gly Asp Ile Asn Asn Leu Ala Phe Ile Asn Cys Ile Asn  
1569 180 185 190

pp 4-6

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/786,880B

DATE: 04/12/2002

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

```

1571 Tyr Ser Arg Ser Gln Leu Phe Asn Ala His Glu Leu Leu Gly Gln Val
1572      195      200      205
1574 Val Phe Gly Leu Ala Asp Asn Tyr Tyr Glu Ser Tyr Gly Thr Leu Asn
1575      210      215      220
1577 Asn Tyr Asn Ser Leu Val Glu Phe Ile Leu Lys Asn Ile Ser Asp Glu
1578 225      230      235      240
1580 Asp Val Phe Val Ile His Phe Leu Pro Ser Thr Leu Gln Leu Phe Lys
1581      245      250      255
1583 Lys Leu Leu Gln Leu Gly Glu Glu Ser Leu Val Asp Gln Phe Tyr Lys
1584      260      265      270
1586 Thr Ile Thr Ser Ser Ile Leu Lys Asp Tyr Glu Ala Asn Asn Phe Ser
1587      275      280      285
1589 Lys Ser Glu Asp Ile Asp Leu Ser Lys Ser Lys Leu Ser Gly Phe Glu
1590      290      295      300
1592 Ile Val Thr Ser Phe Ile Phe Leu Thr Glu Phe Ile Pro Trp Cys Lys
1593 305      310      315      320
1595 Gln Leu Ser Ser Arg Thr Ala Lys Tyr Asp Phe Lys Asp Asp Ile Leu
1596      325      330      335
1598 Lys Tyr Met Glu Phe Leu Ile Ser Tyr Gly Val Met Glu Arg Leu Leu
1599      340      345      350
1601 Ser Tyr Cys Ser Glu Thr Ser Asn Ala Lys Thr Gln Gln Val Tyr Asp
1602      355      360      365
1604 Trp Ser Asn Met Tyr Asp Phe Arg Ala Leu Leu Gln Lys Asn Phe Pro
1605      370      375      380
1607 Arg Leu Thr Pro Ala Lys Phe His Tyr Pro Gly Asn Gln Glu Leu Leu
1608 385      390      395      400
1610 Asn Ala Val Arg Pro Gly Tyr Glu Asn Ile Ser Lys Leu Ile Asp Ile
1611      405      410      415
1613 Ser Phe Leu Thr Leu Asp Pro Ser Leu Asn Glu Thr Leu Val Ser Pro
1614      420      425      430
1616 Phe Phe Gln Ser Phe Phe Ser Val Phe Ile Ser Asn Ala Ala Val Val
1617      435      440      445
1619 Met Thr Ser Leu Arg Asp Ser Glu Glu Asp Phe Val Leu Ser Ser Leu
1620      450      455      460
1622 Asn Glu Ser Asp Glu Glu Glu Glu Glu Glu Ser Asp Ser Asp Glu
1623 465      470      475      480
1625 Asp Ser Ser Thr Pro Lys Asn Lys Glu Lys Ser Ala Gly Leu Asp Leu
1626      485      490      495
1628 Asp Lys Ile Ala Gln Arg Ala Glu Leu Glu Arg Phe Tyr Leu Ala Phe
1629      500      505      510
1631 Ala Tyr Thr Tyr Asn Asn Arg Pro Glu Leu Cys Ala Leu Phe Trp Gly
1632      515      520      525
1634 Asn Glu Gln Val Thr His Asp Ile Ile Gly Phe Ile Ser Trp Gly Leu
1635      530      535      540
1637 Ala Asn Asn Thr Ser Pro Leu Ile Thr Ala Thr Phe Cys Leu Leu Leu
1638 545      550      555      560
1640 Gly Ser Leu Ala Ser Ala Gly Ala Glu Ala Thr Ser Arg Ile Trp Glu
1641      565      570      575
1643 Ile Leu Val His Asn Asn Asn Asn Ala Ser Thr Arg Lys Asn Asp Phe

```

## RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

1644		580		585		590	
1646	Ser	Lys	Val	Ser	Val	Asp	Ser
1647		595		600		605	
1649	Asp	Ser	Leu	Asn	Glu	Ser	Phe
1650		610		615		620	
1652	Leu	Asn	Gln	Lys	Lys	Gln	Asp
1653	625			630		635	
1655	Gln	Asp	Leu	Asp	Asp	Ser	Gly
1656				645		650	
1658	Glu	Asp	Ser	Leu	Val	Leu	Ile
1659				660		665	
1661	Ile	Val	Lys	Asn	Leu	Asn	Thr
1662				675		680	
1664	Ser	Val	Val	Tyr	Thr	Arg	Phe
1665				690		695	
1667	Phe	Asp	Asn	Leu	Ile	Asn	Gly
1668	705					710	
1670	Ile	Gln	Ser	Thr	Asn	Asn	Pro
1671				725		730	
1673	Val	Ser	Asp	Asp	Ser	Arg	Ile
1674				740		745	
1676	Leu	Gly	Asp	Phe	Val	Thr	Asn
1677				755		760	
1679	Ile	Trp	Arg	Leu	Val	Asp	Arg
1680				770		775	
1682	Pro	Glu	Asp	Lys	Lys	Asp	Ala
1683	785					790	
1685	Asn	Ser	Lys	Lys	Asn	Val	Pro
1686				805		810	
1688	Thr	His	Leu	Ser	Gln	Ile	Gly
1689				820		825	
1691	Leu	Thr	Pro	Tyr	Ala	Asp	Ser
1692				835		840	
1694	Leu	Tyr	Pro	Cys	Asp	Leu	Gly
1695				850		855	
1697	Gly	Ile	Trp	Pro	Tyr	Ile	Glu
1698	865					870	
1700	Ser	Gly	Thr	Ile	Ala	Asn	Lys
1701				885		890	
1703	Leu	Leu	Glu	Leu	Phe	Ser	Asn
1704				900		905	
1706	Leu	Ile	Asp	Val	Ala	Pro	Lys
1707				915		920	
1709	Gly	Ile	Phe	Asp	Ser	Leu	Ile
1710				930		935	
1712	Phe	Val	Lys	Leu	His	His	Ser
1713	945					950	
1715	Asn	Arg	Thr	Phe	Ser	Ala	Leu
1716						965	

## RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

```

1718 Ser Val Asn Glu Ser Gly Glu Ser Ala Glu Leu Val Ser His Ala Leu
1719          980          985          990
1721 Gly Leu Ile Asn Ser Leu Leu Arg Val Gln Asn Ser Phe Ile Asn Lys
1722          995          1000          1005
1724 Leu Leu Pro Ile Leu Arg Asn Lys Asp Thr Gln Gln Leu His Arg
1725          1010          1015          1020
1727 Gly Thr Ala Ile Gly Ile Gly Thr Ser Met Ser Leu Ala Leu Ala Thr
E--> 1728 (1025) 1025          1030          1035          1040
1730 Pro Arg Thr Ile Phe Asp Cys Ile Tyr Tyr Pro Lys Asn Leu Gly Thr
1731          1045          1050          1055
1733 His Gly Val Ala Asp Phe Tyr Glu Val Ile Leu Phe His Leu Ser Ala
1734          1060          1065          1070
1736 Val Val Gln Phe Ala Leu Tyr Val Ser Cys Glu Asn Thr Ile Ser Asn
1737          1075          1080          1085
1739 Lys Ala Ile Ser Ile Leu Lys Gly Val Ser Gln Ser Lys Phe Phe Val
1740          1090          1095          1100
1742 Thr Arg Val Ser Ser Ser Ala Asp Pro Leu Leu Asn Asn Asp Arg Leu
E--> 1743 (105) 1105          1110          1115          1120
1745 Ile Thr Thr Phe Glu Asn Ile Asp Glu Ser Ile Lys Ile Lys Phe Ala
1746          1125          1130          1135
1748 Phe Ile Asp Lys Phe Glu Glu Leu Glu Asp Ser Leu Asn Met Lys Tyr
1749          1140          1145          1150
1751 Glu Ile Leu Asp Phe Val Leu Gly Asn Leu Asn Gln Phe Asp Gly Lys
1752          1155          1160          1165
1754 Val Ala Thr Thr Ala His Phe Leu Leu Gly Tyr Lys Val Lys Gly Asp
1755          1170          1175          1180
1757 Thr Leu Asp Leu Val Gln Thr Asn Asp Gln Asn Thr Leu Leu Lys Ser
E--> 1758 (185) 1185          1190          1195          1200
1760 Phe Leu Asn Thr Leu Ser Ile Ser Leu Asp Leu Ile Ser Glu Ile Asp
1761          1205          1210          1215
1763 Tyr Asn Asn Gly Asn Asn His Ile Ile Asp Val Gly Pro Ala Lys Leu
1764          1220          1225          1230
1766 Ser Ser Leu Ile Leu Gln Ile Leu Ile Lys Leu Cys Gln Asp Pro Ile
1767          1235          1240          1245
1769 Ser Ser Ser Ile Thr Leu Asn Gln Leu Arg Glu Tyr Glu Glu Leu Phe
1770          1250          1255          1260
1772 Glu Lys Leu Val Asn Cys Gln Pro Lys Leu Asp Leu Asn Thr Val Trp
E--> 1773 (265) 1265          1270          1275          1280
1775 Cys Gly Asn Gln Phe Asp Gly Asp Leu Gln Ile Asp Ala Ser Asn Val
1776          1285          1290          1295
1778 Phe Val Asp Asn Gln Ala Ser Thr Gln Ala Phe Phe Ser Phe Ile Asn
1779          1300          1305          1310
1781 Gln Arg Asn Leu Ile Leu Gln Tyr Leu Ser Leu Glu Phe His Ser Val
1782          1315          1320          1325
1784 Lys Ser Arg Thr Lys Arg Glu Tyr Tyr Ser Lys Val Leu Thr Asn Asp
1785          1330          1335          1340
1787 Lys Glu Phe Val Asn Arg Thr Pro Lys Val Leu Thr Phe Leu Asn Ile
E--> 1788 (345) 1345          1350          1355          1360
1790 Leu Asn Tyr Ser Phe Lys Asn Phe Glu Val Gln Lys Tyr Glu Trp Leu

```

When  
numbering the  
first amino  
acid on a line,  
please begin  
number directly  
below first  
letter of the  
amino acid  
e.g. Gly |<sup>S</sup>| Thr  
1025 |<sub>P</sub>|

Please ensure  
a space  
exists between  
last number  
and the next  
amino acid

## RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

```

1791          1365          1370          1375
1793 Asp Gln Lys Phe Asn Val Ser Leu Leu Leu Ala Glu Val Asn Ala Gln
1794          1380          1385          1390
1796 Lys Asn Gly Thr Leu Asp Phe Ser Val Leu Thr Lys Val Phe Arg Leu
1797          1395          1400          1405
1799 Leu Cys Gln Thr Ser Asn Leu Ile Thr Pro Glu Ser Lys Gln Leu Phe
1800          1410          1415          1420
1802 Ala Glu Glu Ile Met Val Glu Gly Ser Lys Ile Ser Asp Phe Val Thr
E--> 1803 425          1430          1435          1440
1805 Lys Tyr Leu Val Ser Thr Asp Leu Lys Asp Val Gln Leu Lys Cys Leu
1806          1445          1450          1455
1808 His Ser Trp Cys Gln Leu Ile Glu Ile Leu Val Thr Asp Ser Gly Ile
1809          1460          1465          1470
1811 Asn Ser Leu Asn Phe Ile Leu Glu Val Leu Gln Val Ile Ile Pro Lys
1812          1475          1480          1485
1814 Ile Asn Asp Tyr Phe Asp Val Asp Ile Leu Phe Ser Glu Glu Met Val
1815          1490          1495          1500
1817 Ser Leu Cys Val Leu Leu Phe Asp Leu Tyr Asp Gln Leu Thr Leu Ala
E--> 1818 505          1510          1515          1520
1820 Asp Arg Lys Gly Glu Asp Phe Ala Leu Gly Ile Glu Arg Leu Ile Pro
1821          1525          1530          1535
1823 Leu Phe Gln Thr Cys Ile Ala Gly Ile Leu Asn Ser Asn Ser Thr Pro
1824          1540          1545          1550
1826 Ser Leu Arg Ser Asp Leu Tyr Val Val Gly Asn Lys Phe Leu Leu Lys
1827          1555          1560          1565
1829 Cys Phe Glu Arg Glu Ser Phe Leu Lys Gln Val Met His Ile Ile Lys
1830          1570          1575          1580
1832 Ser Val Asp Lys Lys Phe Phe Gln Val Ile Cys Asn Asp Ala Ile Tyr
E--> 1833 585          1590          1595          1600
1835 Ser Glu Gly Pro Ser Arg Ile Thr Ser Thr Leu Phe Leu Glu Ser Leu
1836          1605          1610          1615
1838 Val His Leu Gly Thr Leu Val Lys Val Asp Phe Ile Leu Asn Ala Leu
1839          1620          1625          1630
1841 Ile Lys Asn Asn Ala Leu Leu Leu Leu Val Arg Ser Val Lys Arg Thr
1842          1635          1640          1645
1844 Asp Ala Met Ile Lys Leu Cys Gln Glu Lys Asn Ser Gly Val Thr Leu
1845          1650          1655          1660
1847 Asp His Phe Ile Phe Asp Leu Met Ala Phe Lys Ala Thr Leu Tyr Phe
E--> 1848 665          1670          1675          1680
1850 Phe Val Arg Val Ala Lys Ser Lys Asn Gly Ala Leu Gln Leu Ile Gln
1851          1685          1690          1695
1853 Asn Glu Leu Phe Ser Ile Leu His Gln Ser Lys Phe Leu Gln Ile Asp
1854          1700          1705          1710
1856 Pro Asp Ile Gly Leu Ser Leu Arg Ile Glu Glu Val Gln Asp His Lys
1857          1715          1720          1725
1859 Thr Val Asn Val Asn Val Leu Leu Asp Thr Pro Leu Ser Ile Thr Asp
1860          1730          1735          1740
1862 Leu Val Asp Pro Tyr Lys Leu Arg Ser Glu Asn Thr Ile Ser Tyr Phe
E--> 1863 745          1750          1755          1760

```

*same  
error*

## RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

1865 Glu Phe Leu Val Pro Ile Phe Gln Leu Leu Thr Thr Val Leu Leu Ser  
 1866 1765 1770 1775  
 1868 Met Gly Pro Asn Tyr Gln Pro Ala Ile Ile Gln Thr Arg Glu Leu Met  
 1869 1780 1785 1790  
 1871 Lys Ser Val Asn Arg Leu Val Val Gly Val Met Lys Arg Asp Phe Leu  
 1872 1795 1800 1805  
 1874 Val Glu Thr Lys Gln Ile Gly Gln Gly Leu Tyr Lys Glu Glu Ser His  
 1875 1810 1815 1820  
 1877 Glu Leu Val Ser Leu Lys Glu Leu Val Lys Leu Phe Ile Leu Ile Asp  
 E--> 1878 825 1830 1835 1840 *same*  
 1880 Ser Leu Ala His Tyr Ser Val  
 1881 1845  
 3121 <210> SEQ ID NO: 21  
 3122 <211> LENGTH: 1257  
 3123 <212> TYPE: PRT  
 3124 <213> ORGANISM: Candida albicans *pp 8-9*  
 3126 <400> SEQUENCE: 21  
 3127 Met Asp Ile Pro Pro Lys Pro Thr Leu Lys Ala Ile Lys Lys Phe Arg  
 3128 1 5 10 15  
 3130 Thr Leu Asp Glu Ile Lys Tyr Ala Met Lys His Val Phe Gln Asp Ala  
 3131 20 25 30  
 3133 Gln Leu Gly Leu Ala Gly His Arg Lys Leu Val Val Ile Leu Lys Asn  
 3134 35 40 45  
 3136 Val Phe Lys Lys Ala Ile Glu Leu Asn Gln Ile Asn Phe Phe Ala Met  
 3137 50 55 60  
 3139 Cys Phe Thr Lys Leu Leu Ser Lys Val Leu Pro Leu Lys Arg Gly Val  
 3140 65 70 75 80  
 3142 Leu Ala Gly Asp Arg Ile Val Lys Phe Cys Tyr Leu Phe Val Asn Gly  
 3143 85 90 95  
 3145 Leu Val Lys Asp Ala Asn Glu Glu Lys Arg Ser Lys Glu Glu Glu Lys  
 3146 100 105 110  
 3148 Glu Glu Lys Asp Lys Asp Glu Asp Lys Asp Thr Asn Glu Ser Asp Lys  
 3149 115 120 125  
 3151 Asn Glu Glu Asp Gln Glu Asp Gln Glu Gly Glu Gly Asp Gln Glu Thr  
 3152 130 135 140  
 3154 Pro Ile Ser Glu Phe Ile Ser Tyr Leu Ile Lys Tyr Leu Leu Ser Gly  
 3155 145 150 155 160  
 3157 Ile Glu Ala Lys Asp Lys Leu Val Arg Tyr Arg Val Val Gln Thr Leu  
 3158 165 170 175  
 3160 Ala Tyr Leu Val Glu Phe Leu Thr Glu Ile His Glu Asn Asn Thr Leu  
 3161 180 185 190  
 3163 Glu Ala Leu Tyr Thr Leu Leu Ser Asn Arg Leu Gln Asp Lys Glu Leu  
 3164 195 200 205  
 3166 Ser Ile Arg Ile Gln Ala Val Ala Leu Ser His Phe Gln Leu Phe  
 3167 210 215 220  
 3169 Glu Phe Ser Ile Glu Gly Asp Thr Gly Glu Phe Glu Asp Glu Leu Ile  
 3170 225 230 235 240  
 3172 Ser Ser Asn Gln Ile Gln Asn Lys Leu Ile Asn Ser Ile Gln Asn Asp  
 3173 245 250 255

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/786,880B

DATE: 04/12/2002

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

```

3175 Asp Ser Pro Glu Val Arg Arg Ala Ala Leu Met Asn Leu Val Lys Thr
3176                260                265                270
3178 Gln Asp Thr Ile Pro Ile Leu Leu Glu Arg Ala Arg Asp Ser Asn Ser
3179                275                280                285
3181 Ile Asn Arg Arg Leu Val Tyr Ser Lys Ile Ala Arg Glu Leu Ile Thr
3182                290                295                300
3184 Asp Ser Leu Asp Asp Leu Glu Phe Glu Asp Arg Glu Phe Leu Leu Lys Trp
3185 305                310                315                320
3187 Gly Leu Asn Asp Arg Asp Glu Thr Val Lys Ala Ala Ala Thr Lys Met
3188                325                330                335
3190 Leu Thr Ile Tyr Trp Tyr Gln Ser Val Asn Glu Asp Leu Leu Glu Leu
3191                340                345                350
3193 Ile Asp Gln Leu Asn Val Lys Ser Ala Ile Ala Glu Gln Ala Ile Leu
3194                355                360                365
3196 Ala Phe Phe Lys Asn Lys Pro Glu Val Leu Glu Thr Ile Lys Ile Asp
3197                370                375                380
3199 Glu Ser Tyr Trp Lys Asn Leu Thr Thr Glu Lys Ala Phe Leu Met Arg
3200 385                390                395                400
3202 Thr Phe Tyr Gln Tyr Cys Asn Glu Asn Gln Leu His Ala Leu Met Asp
3203                405                410                415
3205 Ala Asn Phe Pro Glu Leu Leu Asp Leu Ser Ile Thr Leu Glu Lys Tyr
3206                420                425                430
3208 Leu Ser Val Arg Leu Lys Thr Ile Asn Glu Asn Glu Asn Leu Val Lys
3209                435                440                445
3211 Thr Trp Glu Thr Tyr Asn Ala Lys Ile Asp Glu Leu Asp Asp Gln Ile
3212                450                455                460
3214 Phe Ser Leu Glu Asn Gln Ile Ser Arg Ile Asn Thr Asp Ala Asp Asn
3215 465                470                475                480
3217 Phe Arg Lys Ser Leu Ser Asn Ile Glu Glu Asp Ile Ile Glu Ile Asn
3218                485                490                495
3220 Ile Ala Lys Asp Leu Phe Lys Lys Arg Ile Lys Gln Leu Lys Asn Asn
3221                500                505                510
3223 Ser Gly Asn Leu Glu Asp Leu Ile Thr Glu Glu Asn Gln Glu Ile Ala
3224                515                520                525
3226 Asp Gln Ile Lys Asp Phe Leu Met Glu Asp Leu Gln Gln Gln Leu Glu
3227                530                535                540
3229 Asp Ile Asn Lys Asn Leu Asp Glu Ile Glu His His Pro Glu Asp Ile
3230 545                550                555                560
3232 Thr Ala Lys Leu Glu Glu Leu Gln Thr Lys Tyr Asp Ser Cys Ile Arg
3233                565                570                575
3235 Ala Leu Glu Thr Thr Ser Glu Leu Lys Ile Gln Thr Val Gln Ile Phe
3236                580                585                590
3238 Glu Gln Glu His Glu Asn Asp Cys Ile Pro Phe Val Asp Ala Leu Lys
3239                595                600                605
3241 Glu Leu Glu Phe Ile Ile Asn Gln Leu Leu Leu Ile Val Lys Asp Phe
3242                610                615                620
3244 Asp Tyr Gly Asp Glu Met Ala Arg Arg Lys Leu Leu His Ile Ile Arg
3245 625                630                635                640
3247 Met Thr Leu Thr Glu Asp Lys Leu Pro Asp Ala Leu Ile Ser Val Ala

```



## RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

```

3248          645          650          655
3250 Leu Arg Val Leu Arg Ala Leu Ser Ile Asn Glu Lys Asp Phe Val Ser
3251          660          665          670
3253 Met Ala Val Glu Ile Ile Thr Asp Ile Arg Asp Ser Arg Asp Asp Glu
3254          675          680          685
3256 Glu Phe His Ser Ala Ala Ala Thr Phe Asp Asp Asp Asp Asp Ile
3257          690          695          700
3259 Leu Gly Asn Gly Asp Asp Glu Ser Gln Gln Ser Ser Ser Leu Ser Ala
3260 705          710          715          720
3262 Val Thr Lys Lys Arg Arg Ile Glu Pro Asp Met Pro Pro Asp Asp Ile
3263          725          730          735
3265 Val Leu Arg Cys Leu Thr Met Thr Gln Tyr Val Leu Glu Val Ile Thr
3266          740          745          750
3268 His Ser Leu Asp Asp His Leu Ser Leu Ser Ser Ile Tyr Ser Gly Ile
3269          755          760          765
3271 Val Asn Tyr Ala Ile Gln Asn Glu Ser Lys Lys Lys Leu Tyr Leu Ala
3272          770          775          780
3274 Gly Leu Thr Cys Leu Gly Leu Tyr Ser Leu Ile Asp Ser Lys Ile Ala
3275 785          790          795          800
3277 Arg Ile Ala Thr Thr Thr Leu Leu Leu Ala Met Arg Ser Asn Gly Glu
3278          805          810          815
3280 Glu Val Lys Glu Ile Gly Met Lys Ala Ile Val Asp Ile Leu Ala Ile
3281          820          825          830
3283 Tyr Gly Met Ser Ile Leu Asp Lys Ser Ser Lys Tyr Lys Tyr Ser Arg
3284          835          840          845
3286 Met Phe Phe Lys Val Leu Asn Ser Phe Asp Ala Pro Lys Leu Gln Cys
3287          850          855          860
3289 Ile Val Ala Glu Gly Leu Cys Lys Leu Phe Leu Ala Asp Ile Leu Tyr
3290 865          870          875          880
3292 Lys Thr Asp Lys Arg Ser Leu Phe Gly Asn Ala Ile Gln Gly Gly Gly
3293          885          890          895
3295 Gly Gly Gly Gly Gly Asn Asp Asp Pro Thr Thr Thr Asn Asp Asp Glu
3296          900          905          910
3298 Thr Glu Glu Glu Thr Asp Arg Glu His Glu Lys His Leu Phe Glu Ala
3299          915          920          925
3301 Ile Val Leu Ile Tyr Phe Asn Pro Asn Thr Lys Ser Asn Gln Glu Leu
3302          930          935          940
3304 Gln Gln Ile Leu Ser Phe Cys Ile Pro Val Tyr Ala Phe Ser His Ile
3305 945          950          955          960
3307 Asn His Gln Ile Asn Leu Ala Ala Val Ser Gly Asp Val Ile Tyr Arg
3308          965          970          975
3310 Leu Phe Thr Glu Thr Glu Thr Glu Leu Ser Pro Ser Val Ile Ile Pro
3311          980          985          990
3313 Gln Leu Ile Ser Trp Cys Asp Pro Arg Asn Leu Val Lys Leu Ser Asn
3314          995          1000          1005
3316 Glu Glu Ile Asn Gln Ala Thr Ser His Leu Trp Gln Cys Val Tyr Leu
3317          1010          1015          1020
3319 Leu Gln Val Val Glu Gln Val Asp Ala Arg Asn Val Lys Arg Cys Ile
E--> 3320 025/025          1030          1035          1040

```

## RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

3322 Ile Asn Asn Leu Asn Lys Phe His Ile Thr Glu Glu Leu Glu Ser Asn  
 3323 1045 1050 1055  
 3325 Gln Leu Gln Ala Leu Ile Lys Ala Leu Asp Ala Thr Val Glu Leu Phe  
 3326 1060 1065 1070  
 3328 Thr Asn Asn Glu Asp Asn Pro Asn Phe Ile Leu Asp Lys Pro Thr Lys  
 3329 1075 1080 1085  
 3331 Lys Asn Phe Asp Thr Phe Ile Glu Ser Ile Lys Asn Lys Leu Glu Ile  
 3332 1090 1095 1100  
 3334 Ala Gln Lys Arg Glu Glu Asn Glu Leu Ile Lys Ser Gly Thr Asn Ser  
 E--> 3335 105 1110 1115 1120 *same*  
 3337 Ile Leu His Glu Leu Asp Asp Leu Asp Ile Gly Thr Gly Glu Ser Ser  
 3338 1125 1130 1135  
 3340 Gln Ile Ser Ile Lys Ser Glu Thr Lys Arg Arg Asp Leu Asp Arg Ser  
 3341 1140 1145 1150  
 3343 Leu Gln Val Ser Lys Thr Thr Ser Pro Glu Thr Ser Glu Asn Glu Asp  
 3344 1155 1160 1165  
 3346 Glu Glu Asp Asp Asn Glu Glu Glu Glu Gln Glu Lys Lys Lys Ser Phe  
 3347 1170 1175 1180  
 3349 Thr Asp Gly Lys Asn Lys Leu Glu Leu Lys Ala Asp Lys Pro Ile Thr  
 E--> 3350 185 1190 1195 1200 *same*  
 3352 Phe Lys Ala Glu Asp Lys Arg Glu Gly Ser Val Glu Thr Asp His Gly  
 3353 1205 1210 1215  
 3355 Gln Glu Gln Val Leu Val Glu Ser Lys Lys Val Ile Asp Ser Asn Val  
 3356 1220 1225 1230  
 3358 Glu Asp Ser Leu Glu Asp Ile Asp Lys Phe Leu Glu Glu Ala Asp Asp  
 3359 1235 1240 1245  
 3361 Val Asp Tyr Gly Asp Ile Ser Met Asp  
 3362 1250 1255  
 3464 <210> SEQ ID NO: 25  
 3465 <211> LENGTH: 231 ? See below  
 3466 <212> TYPE: DNA (over 231 bases shown)  
 3467 <213> ORGANISM: Artificial Sequence *see p. 10, too*  
 3469 <220> FEATURE:  
 3470 <223> OTHER INFORMATION: Description of Artificial Sequence: Homologous  
 3471 Fragment to Sc YOR110  
 3473 <400> SEQUENCE: 25  
 3474 atatgtgttg atagttacac atgcagcaac gaaaattgct ttaggatcag ctttattaca 60  
 3476 gttaaaatca gttactgatg ttatagatga taatcaaact gtgttacgtg ctggtgcatg 120  
 3478 ttcattatcc aaatttgta gagatggcga agataaaacc aatcatacta ttcaatggaa 180  
 3480 aattgtcatg aatggttaatt gtgaattcgt gacacagggt gaagaaatga a  
 E--> 3484 aaa gag att gga atg aaa gct att gtg gat ata ttg gca att 2496  
 3485 Glu Val Lys Glu Ile Gly Met Lys Ala Ile Val Asp Ile Leu Ala Ile 231  
 W--> 3486 820 825 830  
 E--> 3488 tat ggt atg agt att ctt gat aaa tca tct aaa tac aaa tat tca aga 2544  
 3489 Tyr Gly Met Ser Ile Leu Asp Lys Ser Ser Lys Tyr Lys Tyr Ser Arg  
 W--> 3490 835 840 845  
 E--> 3492 atg ttt ttc aaa gtt tta aat tca ttt gat gca cca aaa tta caa tgc 2592  
 3493 Met Phe Phe Lys Val Leu Asn Ser Phe Asp Ala Pro Lys Leu Gln Cys  
 W--> 3494 850 855 860

FyI: Codons (triplets) must appear directly above

their respective amino acids. Do not show

amino acids with no  
codon above them. = 4/12/02

## RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:46

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

E--> 3496 att gtc gct gaa gga tta tgc aaa ttg ttt tta gcc gat att ttg tac 2640  
 3497 ile Val Ala Glu Gly Leu Cys Lys Leu Phe Leu Ala Asp ile Leu Tyr  
 W--> 3498 865 870 875 880  
 E--> 3500 aag act gac aaa cgg agt tta ttt gga aat gct att caa ggt ggt ggt 2688  
 3501 Lys Thr Asp Lys Arg Ser Leu Phe Gly Asn Ala ile Gln Gly Gly Gly  
 W--> 3502 885 890 895  
 E--> 3504 ggt ggt ggt ggt ggt aat gat gat cca act acc acc aat gac gat gaa 2736  
 3505 Gly Gly Gly Gly Gly Asn Asp Asp Pro Thr Thr Asn Asp Asp Glu  
 W--> 3506 900 905 910  
 E--> 3508 act gaa gaa gaa aca gat cga gag cat gaa aag cat tta ttt gaa gcg 2784  
 3509 Thr Glu Glu Glu Thr Asp Arg Glu His Glu Lys His Leu Phe Glu Ala  
 W--> 3510 915 920 925  
 E--> 3512 att gta ctt att tat ttc aac ccc aac acc aaa tca aat caa gaa tta 2832  
 3513 ile Val Leu ile Tyr Phe Asn Pro Asn Thr Lys Ser Asn Gln Glu Leu  
 W--> 3514 930 935 940  
 E--> 3516 caa caa att ttg tca ttt tgt att cca gtt tat gcc ttt tct cat ata 2880  
 3517 Gln Gln ile Leu Ser Phe Cys ile Pro Val Tyr Ala Phe Ser His ile  
 W--> 3518 945 950 955 960  
 E--> 3520 aat cat caa atc aat tta gct gca gtt agt ggt gat gtt att tat cga 2928  
 3521 Asn His Gln ile Asn Leu Ala Ala Val Ser Gly Asp Val ile Tyr Arg  
 W--> 3522 965 970 975  
 E--> 3524 ctt ttc act gaa aca gaa aca gaa tta tca cca agt gtt ata atc cct 2976  
 3525 Leu Phe Thr Glu Thr Glu Thr Glu Leu Ser Pro Ser Val ile ile Pro  
 W--> 3526 980 985 990  
 E--> 3528 caa tta ata tca tgg tgt gat cct cga aat tta gtt aaa tta tgg aat 3024  
 3529 Gln Leu ile Ser Trp Cys Asp Pro Arg Asn Leu Val Lys Leu Ser Asn  
 W--> 3530 995 1000 1005  
 E--> 3532 gag gaa ata aat caa gca aca tca cat tta tgg caa tgt gtt tat tta 3072  
 3533 Glu Glu ile Asn Gln Ala Thr Ser His Leu Trp Gln Cys Val Tyr Leu  
 W--> 3534 1010 1015 1020  
 E--> 3536 tta caa gtg gtt gaa caa gta gat gct cgt aat gtt aaa aga tgc atc 3120  
 3537 Leu Gln Val Val Glu Gln Val Asp Ala Arg Asn Val Lys Arg Cys ile  
 W--> 3538 1025 1030 1035 1040  
 E--> 3540 att aac aat ttg a

nos.  
off

← no cumulative  
base total

## VARIABLE LOCATION SUMMARY

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:47

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

Use of n's or Xaa's (NEW RULES):

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of &lt;220&gt; to &lt;223&gt; is MANDATORY if n's or Xaa's are present.

in &lt;220&gt; to &lt;223&gt; section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:12; N Pos. 338,339,340,341,342,344,347,349,366,368,383,387,407,408,433

## VERIFICATION SUMMARY

DATE: 04/12/2002

PATENT APPLICATION: US/09/786,880B

TIME: 14:58:47

Input Set : A:\446.001.txt

Output Set: N:\CRF3\04122002\I786880B.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:1728 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:11 —  
M:332 Repeated in SeqNo=11  
L:1903 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:12  
L:1903 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:12  
L:1903 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:300  
L:1905 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:12  
L:1905 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:12  
L:1905 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:360  
L:1907 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:12  
L:1907 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:12  
L:1907 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:420  
L:2219 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:16  
L:3320 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:21 ✓  
M:332 Repeated in SeqNo=21  
L:3484 M:254 E: No. of Bases conflict, LENGTH:Input:2496 Counted:273 SEQ:25  
L:3486 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
M:254 Repeated in SeqNo=25  
L:3490 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3494 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3498 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3502 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3506 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3510 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3514 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3518 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3522 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3526 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3530 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3534 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3538 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:25  
L:3540 M:252 E: No. of Seq. differs, <211> LENGTH:Input:231 Found:910 SEQ:25